Which one is me? Identifying Oneself on Public Displays

Mohamed Khamis[1], **Christian Becker**[1], **Andreas Bulling**[2], **Florian Alt**[1,3]
[1] Ubiquitous Interactive Systems Group, LMU Munich, Germany
[2] Max Planck Institute for Informatics, Saarland Informatics Campus, Germany
[3] Munich University of Applied Sciences, Germany
mohamed.khamis@ifi.lmu.de, bulling@mpi-inf.mpg.de, florian.alt@ifi.lmu.de

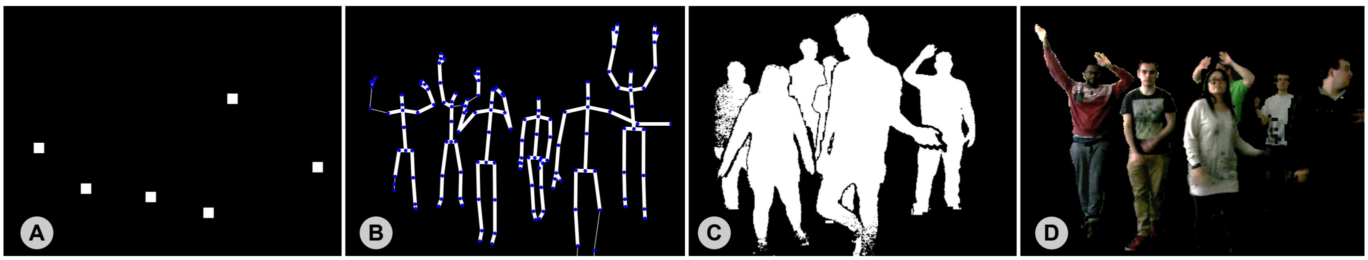

Figure 1. We studied the effect of user representations on users' ability to distinguish their own on-screen representations. We experimented with four user representations: (A) abstract objects, (B) skeletons, (C) silhouettes, and (D) mirror images. In a prestudy, we uncovered five strategies users employ to identify themselves. In a subsequent study we quantified the recognition time and accuracy of each representation with respect to these strategies. We conclude by six recommendations to guide designers in picking the representation most suitable for their deployment's context.

ABSTRACT
While user representations are extensively used on public displays, it remains unclear how well users can recognize their own representation among those of surrounding users. We study the most widely used representations: abstract objects, skeletons, silhouettes and mirrors. In a prestudy (N=12), we identify five strategies that users follow to recognize themselves on public displays. In a second study (N=19), we quantify the users' recognition time and accuracy with respect to each representation type. Our findings suggest that there is a significant effect of (1) the representation type, (2) the strategies performed by users, and (3) the combination of both on recognition time and accuracy. We discuss the suitability of each representation for different settings and provide specific recommendations as to how user representations should be applied in multi-user scenarios. These recommendations guide practitioners and researchers in selecting the representation that optimizes the most for the deployment's requirements, and for the user strategies that are feasible in that environment.

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces—*Graphical user interfaces (GUI), Screen design, User-centered design*

Author Keywords
Public Displays; User Representations; Multiple Users

CHI 2018, April 21–26, 2018, Montreal, QC, Canada.

ACM 978-1-4503-5620-6/18/04...$15.00.
https://doi.org/10.1145/3173574.3173861

INTRODUCTION
Interactive displays have become ubiquitous in public spaces such as museums, airports, and shopping malls. In the vast majority of cases people approach public displays in groups [7, 11, 26], which led to an increasing number of very large displays to allow interaction by multiple users [1, 2]. Multi-user interactive public displays often respond to individual users by assigning a visual representation to each user [17, 33, 38]. The past years witnessed an extensive employment of user representations on the display. These representations can serve multiple purposes; previous work employed them to attract the attention of passersby [26, 36], initiate interaction with the display [35], and provide real-time feedback to users [1]. In most of these works, multi-user interaction is supported by showing multiple on-screen representations simultaneously. Although previous research investigated different aspects of user representations, to date, an understanding of how quickly and accurately users can discern their own on-screen representation from multiple other ones is still missing.

Previous work highlighted the importance of this problem. For example, Müller et al. reported that when a crowd gathers in front of a display, it becomes difficult to distinguish which effect is caused by whom [26]. Wouters et al. [37] also reported that passersby lose motivation to interact with public displays when they are unable to identify themselves, which is often the case when it is crowded. Bridging this knowledge gap is valuable for the display community, since it supports researchers and practitioners designing interactive display applications to choose the right user representation. In some cases it might be important to minimize interaction times and hence support users in findings their own representation as quickly as possible. This is mainly the case when displays are used as tools [26]. Yet, especially for more playful applications (displays as toys [26]) the challenge of identifying oneself could be a playful element as part of an interactive game. For example, application designers may want to avoid a particular representation, such as a mirror image, because it invades privacy or makes users feel uncomfortable in public. Designers may prefer a representation, for example a silhouette, because it can be integrated with the application content by employing the corporate color.

To be able to provide recommendations as to how user representations should be employed in multi-user scenarios where performance matters / does not matter, we conducted an in-depth investigation of the influence of different user representations on user performance. In particular, we investigated the most commonly employed user representations: abstract objects (e.g., the user is represented by an abstract square), a skeleton (e.g., provided by Microsoft Kinect), a silhouette, and a mirror image of the user (see Figure 1). At the outset of our research, we investigated users' strategies when it comes to identifying themselves among other users (N=12). Such strategies include changing position, adopting a particular body posture, observing the appearance of the representation and observing the behavior of others. Subsequently, we conducted a controlled experiment (N=19) where we quantified accuracy and time needed to identify representations based on different strategies and representations types. We found noticeable differences in self-identification time based on the user representation type, the movement strategies and specific combinations of both, whereas accuracy is in many cases negligible. Our work is complemented by providing specific recommendations as to which representations should be used, given the context of a deployment and the content of the employed application.

The contribution of this work is threefold: (1) We identify *five main strategies* employed by users to distinguish their own representations among multiple on-screen ones through a qualitative user study (N=12). (2) We report on a subsequent quantitative study (N=19) where we measured the *impact of different types of user representations* (Figure 1) *and strategies* on the users' speed and accuracy in identifying their own representation among others. (3) We compile a *set of recommendations* that guide designers in choosing the right representation depending on the context of the deployment.

RELATED WORK

We build on two strands of prior research: Supporting multi-user interaction on public displays, and user representations.

Multi-user Interaction on Public Displays

Public displays are often approached by groups [7]. In cases where only one user is interacting initially, the honeypot effect attracts passersby, eventually leading to multi-user interaction [26]. To address this, public displays increasingly support interaction for multiple users through different modalities. Examples for multi-user interaction using touch include Hello.Wall [29], EyePACT [13] and Worlds of Information [10]. Systems such as Looking Glass [26], Media Ribbon [2], StrikeAPose [34] and MyPosition [31] support multi-user interaction by mid-air gestures. Gaze-enabled public displays have also started to support multiple users [12, 16, 38]. Other displays allow interaction via mobile devices [14, 18, 32].

As users approach a display it is usually not clear to them, whether a public display application supports multiple users. This problem is mitigated in cases where user representations are employed, since in this case it is communicated to passersby that everybody being represented on the display can interact. At the same time, in multi-user scenarios it is often challenging for users to know who caused which effect [26]. Our work contributes in understanding which representations work best in helping public display users identify themselves in multiple users scenario. This knowledge can be exploited by designers of public display applications to support or – if desired – make it difficult for users to identify themselves.

User Representations

Research in psychology shows that humans are highly capable of recognizing themselves and others through motion [6] and mirror images [22]. Prior work investigated different ways of representing users on public displays: *Appearance matching* representations include silhouettes [2] (aka contours [31, 34]) and mirrors [26] (Figures 1C and 1D). *Kinesthetic-visual matching* ones include abstract objects [19, 35, 37], avatars [26], and skeletons [2, 35, 36] (Figures 1A and 1B).

Previous work found a correlation between vagueness of the representation and willingness to vote in public [18, 31], implying that anonymity of the representation is sometimes desired. Ackad et al. [2] found that, compared to silhouettes, skeletons are perceived to be more playful, resulting in longer interaction times [30]. However, users stayed more focused when interacting with silhouettes [2, 36]. Skeletons were used to instruct users how to interact with displays [3], while skeleton-like avatars and silhouettes were utilized to register users at displays to kick off interaction [35]. In terms of communicating interactivity, lab studies showed that mirror representations do not greatly outperform silhouettes [26, 28], but a significant improvement in favor of mirrors was observed in a field study [26]. In contrast to other representations, mirrors were associated with privacy concerns [26].

From this we learn that there seems to be no representation that is generally superior to others and that designers need to decide in-situ which representation to use. Deployment contexts highly influence the behavior of passersby, for example, a display in a hallway would expect walking users, while a display in an elevator or next to a vending machine would expect its users to be stationary. Furthermore, the location may also impact whether or not users care about being represented by their mirror image or a more abstracted representation. In order for designers to make an informed decision, a comprehensive understanding of the effects caused by using such representations is necessary. In particular, knowledge on performance in terms of time and accuracy is missing as of today – yet it is an important aspect when it comes to choosing a specific representation. Based on our findings, we conclude this work with recommendations on how to determine which user representations are appropriate for which scenarios.

INVESTIGATED USER REPRESENTATIONS

Based on related work, we investigate four user representations that cover the full range of possible details representations can provide. Two are appearance matching representations: silhouettes [2, 26] (aka contours [24, 31, 34]) and mirrors [26] (Figures 1C and 1D). And two are kinesthetic-visual matching representations: abstract objects [19, 35, 37] and skeletons [2, 35, 36] (Figures 1A and 1B).

Abstract objects. A distinction of abstract objects, is that they can support a top-view. A top-view is a plausible design for multiuser scenarios; a motion sensing device at the top would minimize the chances of users occluding each other. In this representation, moving forward corresponds to an upward movement of the representation, while a move to the back traverses it downwards (see Figure 1A).

Skeletons. While abstract objects react only to the user's movements in the interaction space, skeletons react to other body movements such as moving the arms and legs. Hence allowing kinesthetic-visual matching [2]. However skeletons also reflect aspects of the user's appearance, such as the user's height and body posture (see Figure 1B).

Silhouettes. Silhouettes reflect more details about the user's appearance such as the body figure, hair appearance and in some cases accessories such as bags and headwear. They are classified as appearance matching representations [2] since users rely mainly on matching their own appearance with that of the Silhouette. However, matching body movements can also help in identifying oneself (see Figure 1C).

Mirrors. In this representation users have the largest variety of cues that can be utilized to identify themselves. In addition to the cues provided by skeletons and silhouettes that are also provided by mirrors, users can match their faces, clothes (e.g., shirt color), hair and skin color (see Figure 1D).

PRESTUDY: SELF-IDENTIFICATION STRATEGIES

Before studying which representation helps users the most in identifying themselves among several ones, it was essential to understand what strategies users employ to achieve that. This was done to ensure that the results of the subsequent study are not biased due to a certain representation–strategy combination that might not always be possible in real deployments. For example, relying on strategies that include walking in front of a display or performing extensive arm movements may not be feasible in narrow areas such as in a hallway [5], on an escalator, or in an elevator. On the other hand, a user could be moving before noticing displays deployed in streets or hallways – hence, it is plausible to exploit this movement.

Setup and Implementation

We deployed a Kinect One in front of a projected display in a $5.6\,\text{m} \times 5\,\text{m}$ room. We chose to conduct the study in a large room to avoid restricting possible strategies. The projection had a diagonal of approximately 2 meters and a resolution of 1280×720 pixels (see Figure 2).

Abstract objects were implemented by mapping the position of the user's sacroiliac joint (at the bottom of the spine) to the screen. The skeleton objects were implemented in a similar manner. However in this case not only the sacroiliac joint was used but all skeletal joints provided by the Kinect API. For each joint a circle was drawn on the screen indicating its spatial position. For silhouettes and mirrors we used the depth stream provided by Kinect to identify the skeleton. This allowed us to identify which pixels in the video stream belong to each user. We generated silhouettes by changing the color of these pixels to white and subtracting the background. Mirror objects were created by combining the 2D-color-stream and the depth stream provided by the Kinect API.

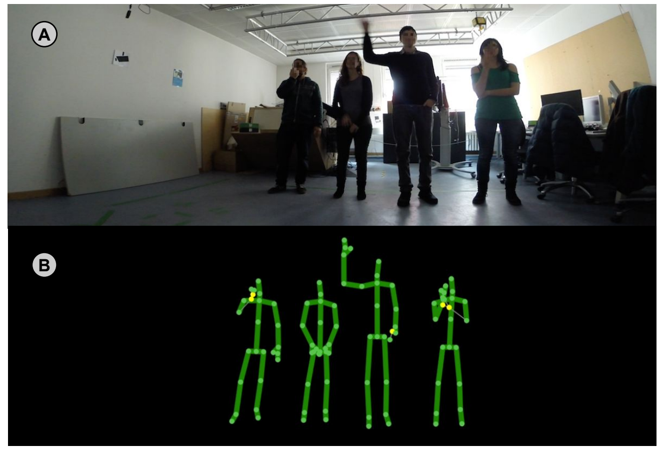

Figure 2. The prestudy participants stood at a distance of 3.6 meters away from the projection. They were then shown one of the four representation types at a time, and asked to identify themselves as quickly as possible. Afterwards the participants were interviewed to uncover the strategies they employed to perform the task.

Study Design and Procedure

We invited 3 groups of 4 participants each (7 females) with ages between 21 and 33 ($M = 25.3$, $SD = 3.4$) through university mailing lists. In a within-subjects repeated measures experiment, with a single variable being the user representation type, participants were asked to stand approximately 3.6 meters away from the projection (see Figure 2A). The experimenter then explained the study and asked the participants to identify their representations as fast and as accurately as possible. No hints were given to the participants regarding how to identify themselves. The experimenter launched an application that showed one type of user representation at a time for all participants (Figure 2B shows the skeleton representation condition). The experimenter traversed through the user representation types. Each representation was used 6 times in a counter-balanced manner. In each run, the experimenter waited until all participants in the current group confirmed that they recognized themselves, and then proceeded to the next representation. With the consent of participants, the sessions were video-recorded for post-hoc analyses. Afterwards, participants were interviewed and asked to fill in a questionnaire.

Findings

Although we foretold that participants will succeed in their task, the idea was to present the participants with a realistic situation where recognizing their representations is possible in order to study their strategies.

Each participant performed 24 runs (6 runs $\times$ 4 representations). We classified the strategies into five main strategies.

Participants' first intuition was to move their arms. When shown the abstract objects, they quickly realized that the representation does not respond to arm movements and started moving around in the room instead. They tried to match their movement trajectory to that of the on-screen objects. Some reported that in their effort to make their movement trajectory stand out, they kept an eye on all on-screen moving objects and tried to move in distinct ways.

Strategy 1 – Motion-based Kinesthetic Visual Matching: Users move in front of the display to help identify their abstract representation.

To recognize their skeleton and silhouette representations, participants either walked only, moved their arms only, or both moved their arms and walked around. Inline with previous work [2, 30], it was noticed and also reported by participants that they find the skeletons to be playful, and hence they spent more time experimenting with different movements. Participants reported that skeletons are easier to recognize compared to silhouettes due to their 3D nature. For example, arms can still be seen when they are in front of the body.

Strategy 2 – Body gestures-based Kinesthetic Visual Matching: For less abstract representations, such as skeletons and silhouettes, users wave their arms, walk around, or walk around while moving their arms.

Participants were able to match their appearance to silhouettes after seeing their silhouette at least once. Before its first appearance, kinesthetic-visual matching was used (Strategies 1 and 2). When asked, participants explained that once they saw how their own silhouette looks like, they were easily able to spot it whenever they saw it again. Mirrors were spotted as soon as they were revealed to the participants; most of the times there was no need to move.

Strategy 3 – Appearance Matching: Appearance matching is always used for mirrors, and also for silhouettes after the users had seen their silhouette at least one time before.

In addition to movements, participants noted that their position relative to the display and to the other participants was influential. For example, P1 noted that knowing she was always depicted by the leftmost representation made it easier for her to find herself the following times.

Strategy 4 – Relative Position Mapping: Users utilize their position relative to the display and to other users to identify their representation.

The interviews also revealed that the participants observed the movements of others around them and tried to stand out by performing unique movements. For example, one participant quickly stepped forward and backward when she noticed the others were moving slowly towards the sides. Moreover, it was noticed that participants who were faster in determining their representations behaved differently after completing their task (e.g. stopped moving, looked towards the experimenter). This was noticeable by other participants and, in turn, helped them exclude the representations that are behaving differently.

Strategy 5 – Tracking Surrounding Users: Users utilize the behavior of real surrounding users to identify their representations.

While strategies 1, 2, and 3 were partially discussed in prior work [1, 2, 3, 30], this study is the first to consider self-recognition in case of multiple users. Consequently, strategies 4 and 5 are novel and could not be extracted from prior work.

MAIN STUDY: EFFECTS OF USER REPRESENTATION

The goal of this study was to quantify the independent impact of the different representations on user performance when trying to identify oneself in front of public displays. We studied the time taken by participants to recognize themselves, in addition to the accuracy of the decision, i.e., whether or not the identified representation is indeed the correct one.

Implications of Prestudy on the Main Study Design

The prestudy showed that there are several factors that could have an influence on the identification of representations. To study the impact of these factors, and to distinguish the impact of the representation type from the impact of the combination of a particular representation and one of these factors (e.g., distinguishing the impact of skeleton representations from the impact of the skeleton representation's position relative to the user), it was necessary to make the following design decisions.

Movement Types

The prestudy showed that user movements are heavily influenced by the type of shown representation. In order to distinguish the impact of the representation from that of the representation-movement combination, it was necessary to control this variable. Hence, based on *strategies 1*, *2*, and *3*, we introduced 4 conditions to the main study:

NoMove No movement

Arms Moving arms while stationary

Walk Walking around

Walk+Arms Walking around *and* moving the arms

Effect of User Position

The position of the prestudy participants influenced their ability to identify themselves (*Strategy 4*). Hence, instead of always showing the representation in a mirrored-position, we introduced a variable with two conditions: (1) mirrored-position representation, and (2) randomly-positioned representation (Figure 3). This was done to separate the impact of the position from that of the representation. In the mirrored-position condition, the correct representation is always positioned directly in front of the user, while in the randomly-positioned one, the representation's starting position is randomly chosen.

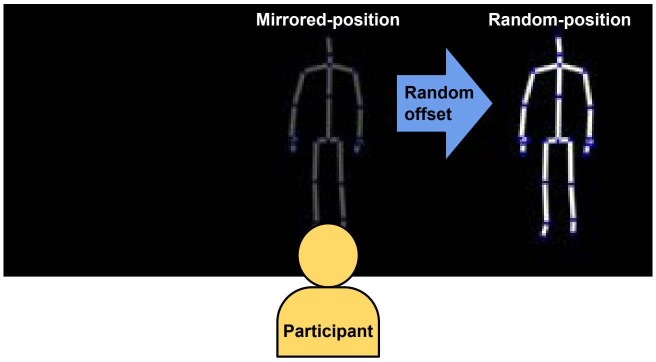

Figure 3. The prestudy showed that displaying the representation in a mirrored-position could influence the user's decision (Strategy 4). Hence we introduced two conditions to the main study in which the representation's starting position is either mirrored or randomly determined at the beginning of the session.

Effect of Surrounding Users

The prestudy participants exploited the behavior (*Strategy 5*) and the position (*Strategy 4*) of those around them to identify their own representations. While this was the case in our prestudy setup, this strategy is not always feasible in real public display deployments. For example, users whose representations are simultaneously shown might not see each other due to physical occlusions (e.g., other users or walls), or might not even share the same physical space like in [4, 9, 25]. To control this but also maintain multiple user representations, we showed fake representations on the screen (see Figure 4).

Main Study Design

The study followed a within subjects design. In addition to the previously mentioned *movement types* and *representation positions*, an additional independent variable in the study was the shown *user representation*. Based on prior work, we chose the abstract objects, skeletons, silhouettes, and mirrors. Each participant performed 96 rounds: 3 blocks with each block covering all 32 conditions (4 movement types $\times$ 2 representation positions $\times$ 4 user representation types). The randomly-positioned representation condition was always used for the first half of each block. This was done to avoid giving the impression that the position of the correct representation is always mirrored. All other conditions were counter balanced using a Latin-square.

Recording Session

As previously mentioned, we opted for using fake on-screen representations along with the participant's real representation due to the aforementioned issues related to the effect of surrounding users (*Strategies 4* and *5*).

To generate realistic fake representations for the study, we invited 15 participants to record their on-screen representations. Out of the 15, six participants were present in front of the Kinect at a time. Each group of six performed each *movement type* for every *user representation*. We made 2 $\times$ two-minute recordings for each movement type and each representation (2 recordings $\times$ 4 representations $\times$ 4 movement types $=$ 32 two-minute recordings). For each recording, we shuffled the participants such that the same group of six never appeared in more than one recording.

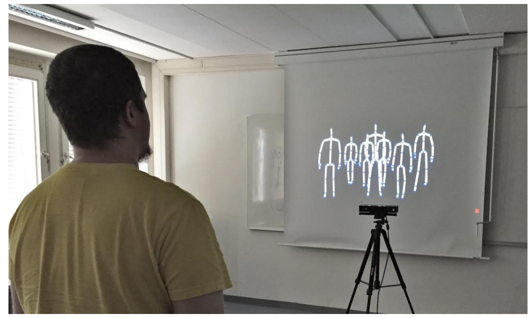

Figure 4. Fake on-screen representations were used along with the participant's real representation. This was done as the presence of multiple real participants would have influenced their decisions due to their behaviors (Strategy 5) and positions relative to each other (Strategy 4).

Participants, Setup and Procedure

We invited 19 participants (8 females) aged between 18 and 36 ($M = 26.5$, $SD = 3.8$), with heights ranging from 156 cm to 206 cm ($M = 1.78$, $SD = 0.12$). None had participated in the prestudy nor in the recording session.

In a setup similar to that used in the prestudy, participants were invited one at a time and were asked to stay within a marked area to ensure they stay within the Kinect's range.

The experimenter started by explaining the study and asking the participant to sign a consent form. The participant was handed an Xbox controller. We logged the timestamps at which any controller button was pressed. This was used to calculate the *recognition time*, which is the time in milliseconds from the moment participants saw the representations until the moment participants pressed a button to signal that they recognized what is thought to be the correct representation. As holding the controller using both hands could have an impact on the recognition of some representations, the participant was asked to hold it with one hand.

After finding the correct representation, the participant verbally communicated the chosen representation to the experimenter (e.g., "It's the third one from the right", "that one, at the bottom left"). The experimenter used a laser pointer to confirm that he understood the participant's choice. This allowed us to calculate the *error rate*, which is a binary value that indicates whether or not the participant's guess was correct.

We preferred not to log the end of the recognition period manually since it is likely it will take additional time from the moment the participant communicates that they identified themselves until the experimenter logs the time. However logging the beginning of the recognition period the way we did was essential to ensure that the participant is performing the movement by the time he/she sees the representations.

At the beginning of each of the 96 rounds, participants were asked to close their eyes. This gave the experimenter the chance to load one of the two recordings that cover the respective condition and launch it. Participants were asked to start performing the respective movement type with eyes-closed, and to keep performing it until the end of the trial; this movement type corresponded to the movement type of the current fake representations. Once participants started, the

Significantly different Representations		$p <$
Abst. objects (6255 ms)	Skeletons (3830 ms)	0.005
Abst. objects (6255 ms)	Silhouettes (2510 ms)	0.001
Abst. objects (6255 ms)	Mirrors (1376 ms)	0.001
Skeletons (3830 ms)	Silhouette (2510 ms)	0.001
Skeletons (3830 ms)	Mirrors (1376 ms)	0.001
Silhouettes (2510 ms)	Mirrors (1376 ms)	0.001
Significantly different Movements		$p <$
NoMove (4499 ms)	Walk (2874 ms)	0.001
NoMove (4499 ms)	Walk+Arms (2522 ms)	0.001
Arms (3468 ms)	Walk+Arms (2522 ms)	0.05

Table 1. This means that mirrors are significantly fastest to recognize, followed by silhouettes, skeletons, and abstract objects respectively. Recognition time is significantly faster in Walk+Arms than in NoMove and Arms, and significantly faster in Walk than in NoMove.

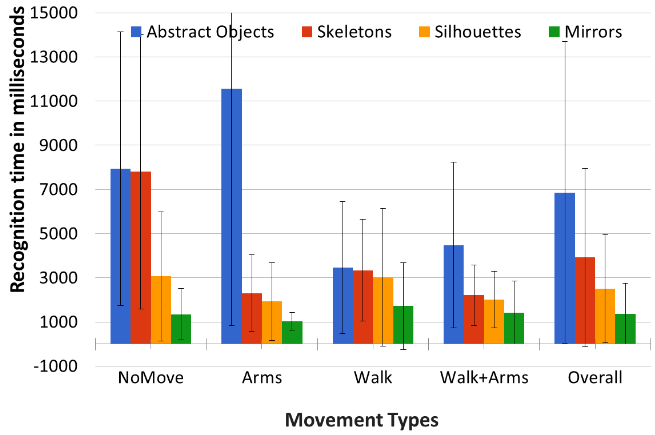

Figure 5. The graph shows the recognition time. Pairwise comparisons are shown in Table 2.

experimenter counted from 1 to 3. Participants were previously instructed to open their eyes at the count of 3, which the system logs as the *start time*. This means that by the time participants' eyes are open, the participants and the fake representations would be all performing the same movement type. This was necessary to avoid influencing participants; for example, if the fake representations started moving before the participant, the participant's stationary representation would be easily noticeable;

Finally, participants were asked to fill in a questionnaire to collect their subjective feedback about the recognition time associated with each representation.

Limitations

As users walk around (Arms and Walk+Arms) silhouette representations might hide each other. This could speed up or slow down recognition times depending on the position of the user. We controlled this effect as much as possible in the main study by defining a starting position. However this effect is expected and could probably have an impact in field deployments.

Behavior in real deployments is expected to be more similar to the prestudy than to the main study, in the main study we aimed to understand the independent effect of each movement/representation type. Thus, it was essential to exclude external influences. For example, had the participants been allowed to stand still while fake skeletons are moving, our results could have mislead the reader into thinking that skeletons are as fast to recognize as mirrors.

A limitation in the design is that we use the same room for all conditions, hence we do not cover different environments (e.g., narrow vs. wide rooms). Nevertheless, the study covers the movement types expected in different environments (e.g., NoMove in narrow areas).

MAIN STUDY RESULTS

Outliers

Previous work showed that users require a minimum of 600 ms for the mental act of routine thinking, and 100 ms for the press/release of a button [15]. Hence we decided to exclude all recognition times that are less than 700 ms (53 out of 1824 time measurements). We expect that in these cases participants either pressed too early by mistake or because they were overly confident that they will recognize their (e.g., mirror) representation immediately.

From inspecting box plots that we drew for each condition, we further excluded the following recognition times which were beyond the quartile by one and a half interquartile range per condition – and not across all of them:

- 26 recognition times for abstract objects ($\geq$ 21249 ms).
- 43 recognition times for skeletons ($\geq$ 9390 ms).
- 33 recognition times for silhouettes ($\geq$ 5651 ms).
- 30 recognition times for mirrors ($\geq$ 2505 ms).

The outliers were caused either by participants forgetting to press the button, or due to spending too much time trying to find their representation in the *no movement* condition. This means that in total we excluded 185 out of 1824 recognition time measurements.

It was necessary to clear out outliers that could bias the results. Not excluding the outliers could wrongfully imply that the benefits of a representation type outweigh its drawbacks, or underrate the average performance of a representation although it performed well in certain conditions.

Recognition Time

For statistical analysis of the recognition time measurements, a linear mixed model was employed due to its robustness against excluded outliers. We then performed post-hoc pairwise comparisons with Bonferroni corrections.

Significant main effects were found for the representation type $F_{3,100.1} = 47.3$, $p < 0.001$ and the movement type $F_{3,61.4} = 10.9$, $p < 0.001$ on the recognition time. There was an interaction between the representation type and the movement type $F_{9,96.1} = 6.28$, $p < 0.001$. This means that the type of shown representation, the type of the user's movement, and the combination of both all influence how fast users detect own representations.

Significantly different Representations		
NoMove: Standing still		$p <$
Abst. objects (7950 ms)	Silhouettes (3063 ms)	0.005
Abst. objects (7950 ms)	Mirrors (1345 ms)	0.05
Silhouettes (3063 ms)	Mirrors (1345 ms)	0.01
Skeletons (7803 ms)	Mirrors (1345 ms)	0.001
Skeletons (7803 ms)	Silhouettes (3063 ms)	0.005
Arms: Moving arms while stationary		$p <$
Abst. objects (11562 ms)	Skeletons (2308 ms)	0.005
Abst. objects (11562 ms)	Silhouettes (1930 ms)	0.005
Abst. objects (11562 ms)	Mirrors (1025 ms)	0.001
Skeletons (2308 ms)	Mirrors (1025 ms)	0.005
Silhouettes (1930 ms)	Mirrors (1025 ms)	0.05
Walk: Walking without moving arms		$p <$
Abst. objects (3465 ms)	Mirrors (1716 ms)	0.05
Skeletons (3336 ms)	Mirrors (1716 ms)	0.001
Silhouettes (3025 ms)	Mirrors (1716 ms)	0.01
Walk+Arms: Walking and moving arms		$p <$
Abst. objects (4477 ms)	Skeleton (2208 ms)	0.005
Abst. objects (4477 ms)	Silhouettes (2008 ms)	0.005
Abst. objects (4477 ms)	Mirrors (1424 ms)	0.001
Skeletons (2208 ms)	Mirrors (1424 ms)	0.05

Table 2. When analyzing each movement type individually, the representation always has a significant effect on recognition time.

No effect was found for the representation position ($p > 0.05$) on recognition time. We also did not find any significant effect of the number of trials done so far for each condition ($p > 0.05$), which means that we did not find any evidence of learning effects.

Post-hoc analysis showed that all representation type pairs are significantly different, while three movement type pairs are significantly different (see Table 1). This means that mirror representations are fastest to recognize, followed by silhouettes, skeletons, and abstract objects respectively. Recognition time is significantly faster when walking while moving the arms (Walk+Arms) compared to both: standing still (NoMove), and standing still while moving the arms (Arms). While walking around without moving the arms (Walk) results in significantly faster recognition time compared to standing still (NoMove). The pair of conditions in which the user walks (Walk and Walk+Arms), as well as the pair in which the user is stationary (NoMove and Arms) are not significantly different.

In addition to the overall comparisons of movement types and representation types, we investigated in-depth the effect of the representation type on recognition time with respect to each individual movement type.

After analyzing the effect of the representation type on recognition time for each movement type, we found that it had a significant main effect in NoMove $F_{3,40.9} = 22.15$, $p < 0.001$, Arms $F_{3,30.44} = 18.65$, $p < 0.001$, Walk $F_{3,59.57} = 11.15$, $p < 0.001$ and Walk+Arms $F_{3,48.42} = 9.68$, $p < 0.001$. Post-hoc analysis showed significant differences between multiple pairs (see Table 2).

The results show that:

- **NoMove:** abstract objects and skeletons are recognizable in case of mirrored-position, but they are the slowest to recognize. While mirrors are the fastest to recognize.
- **Arms:** abstract objects are the slowest to recognize, and mirrors are fastest to recognize. However skeletons and silhouette perform only slightly different (see Figure 5), which explains the lack of significant differences between them. An interesting result is that when abstract objects were shown, participants performed slower in this condition compared to the NoMove condition. This can be explained by previous work, which showed that arm movements induce a higher cognitive load [27], which in turn could have slowed participants. We also expect that participants were trying to spot any subtle changes (which never happened), resulting in them waiting longer before making a decision. Nevertheless we cannot generalize this assumption due to the lack of significant differences ($p > 0.05$).
- **Walk:** the performance of abstract objects, skeletons, and silhouette is almost similar in case of the Walk condition, but mirrors perform better than all of them.
- **Walk+Arms:** arm movements during the Walk+Arms condition seem to have a negative impact on performance in the case of abstract objects (compared to Walk). Similar to NoMove vs Arms, we believe that the arm movements distracted the participants when recognizing abstract objects. It is worth noting that this is the only movement type in which mirrors do not perform significantly better than silhouettes.

Participants' feedback (5-point Likert scale; 1=very fast; 5=very slow) collected through the questionnaire support the quantitative results. There is a statistically significant difference in perceived recognition time depending on which representation type was shown, $\chi^2(3) = 52.04, p < 0.001$. Post-hoc analyses using Wilcoxon signed-rank tests with Bonferroni corrections showed that all pairs are significantly different ($p < 0.005$). Participants indicated that recognition time was fastest in the case of mirrors ($M = 1$, $SD = 0$). Second is silhouettes ($M = 1.84$, $SD = 0.9$), followed by skeletons ($M = 2.47$, $SD = 0.7$), then abstract objects ($M = 4.21$, $SD = 0.85$).

Error Rate

We used repeated measures ANOVA (Greenhouse-Geisser corrected) to analyze the data. Significant main effects were found for representation type on error rate $F_{1.3,23.4} = 40.93$, $p < 0.001$. Significant main effects were also found for movement type on error rate $F_{1.8,32.2} = 54.78$, $p < 0.001$. Table 3 summarizes the results of post-hoc analyses using Bonferroni correction, which revealed significant differences between multiple pairs.

The results show that in the case of abstract objects, users make significantly more errors compared to the other representations. On the other hand, there were no errors at all in the case of mirror representations. There are significantly more

Significantly different Representations		$p <$
Abst. objects (31.8%)	Skeletons (7.68%)	0.001
Abst. objects (31.8%)	Silhouettes (3.73%)	0.001
Abst. objects (31.8%)	Mirrors (0%)	0.001
Skeletons (7.68%)	Mirrors (0%)	0.005
Silhouettes (3.73%)	Mirrors (0%)	0.01
Significantly different Movements		$p <$
NoMove (25%)	Arms (13.82%)	0.001
NoMove (25%)	Walk (2.19%)	0.001
NoMove (25%)	Walk+Arms (1.54%)	0.001
Arms (13.82%)	Walk (2.19%)	0.001
Arms (13.82%)	Walk+Arms (1.54%)	0.001

Table 3. Numbers between brackets denote the percentage of entries where participants picked wrong representation. Mirrors are the least error prone, while abstract objects are the most error prone. Error rate is significantly higher in NoMove than in Arms, Walk, and Walk+Arms, and significantly higher in Arms than in Walk and Walk+Arms.

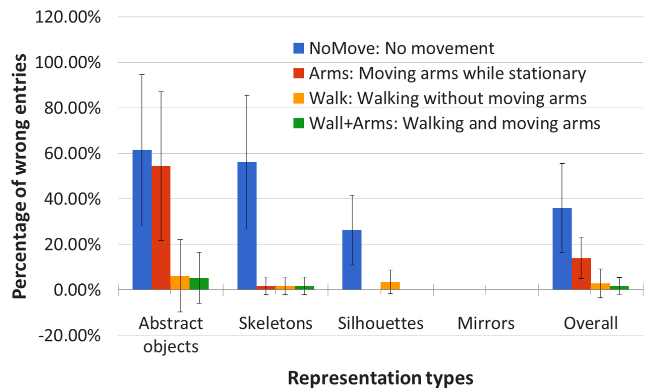

Figure 6. The graph shows the error rate. In the case of abstract objects, users make significantly more errors compared to the other representations. The graph shows that most errors occur in the case of NoMove. Abstract objects are also hard to recognize in Arms. None of the participants performed any errors when selecting from mirror representations.

errors when performing NoMove compared to other movement types, followed by Arms. Figure 6 shows that most of the errors occurred in the case of showing abstract objects and performing NoMove. Many errors also occurred when showing skeletons and performing NoMove.

No significant main effects were found for the representation position on the error rate ($p > 0.05$).

Qualitative Feedback

Participants' feedback focused on the comparison of representations. The vast majority agreed that mirrors are the fastest to recognize. Many reported making use of their clothes (P10, P11, P12, P16) and items such as handbags (P18). For example, P11 noticed that she was the only one wearing blue clothes. P13 on the other hand reported he was sometimes confused in the mirrors conditions because "colors of the clothes are often very similar". Some found silhouettes to be the second best due to the easiness of recognizing "*own body's characteristics (big ears in my case)*" (P3) and the body height (P11). Similar to the feedback from the prestudy, P18 mentioned that he found it easier to recognize his silhouette after "*one or two encounters*". Skeletons also received positive feedback, noting that although it "*has less distinguishing features compared to silhouettes*" (P4), "*at least the movements are the same*" (P6) indicating that movements are perceived to be equally helpful with silhouettes and skeletons. P7 ranked it third because "*even with little movements its still recognizable*". P3 noted that the body position helped decide which skeleton is hers. P13 reported that he sometimes had to guess in the case of abstract objects. All participants found abstract objects to be the most difficult to recognize. P6 noted that it might get even more chaotic with larger crowds.

Apart from recognition, participants also pointed out social and privacy concerns. P13 preferred silhouettes over mirrors because "*a silhouette is less embarrassing in public*". P15 argued that skeletons achieve a balance between recognizability and anonymization, noting that he found it "*scary to see [his] reflection embedded like that*".

DISCUSSION AND RECOMMENDATIONS

In this work we focus on the users ability to recognize their representation, which is a problem at the outset of interaction in a multiuser scenario. The significance of this problem was highlighted in previous research [26, 37].

Our findings confirm several aspects of previous work in the multiuser scenario. For example, in-line with findings by Müller et al. [26], our participants were skeptical about the use of mirrors in public due to privacy concerns.

Privacy Concerns and Potential Implications

Participants who raised privacy concerns were concerned about anonymity; P15 expressed that he was uncomfortable with mirrors, citing how they fail at preserving anonymity. Participants reported that they assume everything is being recorded if they appear on a camera in public space. Since they do not know what the footage is used for, they are often concerned about its misuse. For example, users might think the space owner is sharing their behavioral patterns with third parties without their consent (e.g., to show targeted ads).

A direction for future work is to investigate how these concerns could be influenced by cultural aspects. For example, in some countries CCTV cameras are common, while in others they are considered a breach of privacy. A cross-cultural study could inform the choice of representation depending on the culture the display is deployed in. The user's background can be estimated based on the gestures, language or eye behavior. Silhouettes can then be used instead in such countries.

Social Embarrassment

Some participants expressed concerns about social embarrassment when using mirrors. Observations and feedback from the prestudy confirm that participants find skeletons playful [2] and they extend interaction times [30] beyond the actual task, while silhouettes encouraged serious attention to the task [36].

Recognition Performance

In addition to confirming multiple findings from previous work, results of our main study contribute to better and deeper understanding of how users behave to identify themselves, and how their behavior impacts recognition performance. Namely, the main study shows that *strategies 1*, *2*, and *3* have an impact on recognition performance. Although *strategy 3* also indicated that the perceived recognition time of silhouette representations becomes faster as users become familiar with it, we could not find any significant learning effects in the main study. Furthermore, we could not find a significant effect of the representation position condition that was inspired by *strategy 4* on recognition time and accuracy.

Abstract objects result in significantly slower recognition times compared to other representations in movements that involve moving the arms (Arms and Walk+Arms). Previous work showed that arm movements induce a higher cognitive load [27], which in turn could have slowed participants. We also expect that moving arms in front of a representation that does not respond to arm movements distracts the user and results in longer recognition times (see Figure 5).

Skeletons perform better than abstract objects and silhouettes perform only slightly better than skeletons in all movement types except NoMove (no movement), where it performed significantly better. Although participants of the prestudy noticed that skeletons reflect the 3D position of the arm (e.g., arms are visible even when in front of the user), and hence reported them to be easier to recognize than silhouettes. The main study did not show any significant differences between silhouettes and skeletons in cases where participants moved their arms (Arms and Walk+Arms). This means that in scenarios where these movement types are expected, a designer's choice of silhouettes or skeletons should not be based on recognition time and accuracy, but rather on aspects such as interaction time [30] and attention to the task [36].

Our findings indicate that mirrors outperform other representation types in terms of recognition time and accuracy in every type of movement. This means they are quickly noticeable and thus explains why previous work found that they are significantly better than other representations in communicating interaction to the public [26]. However feedback from our participants shows that they (1) are sometimes socially embarrassing in public, and (2) raise privacy concerns. Additionally, user representations were manipulated in previous work to reflect corporate logo colors or carry content such as buttons and interactive objects [8, 36]. While other user representations are still expected to serve their function after integrating such design changes, mirrors on the other hand would lose core features that make them distinguishable, making them inflexible for these use cases. For example, a user who relies on his shirt's color to identify his mirror representation would be confused if the representation's color was altered, and added content can obscure distinctive parts of the user's appearance.

Recommendation 1: Unless their use contradicts with the display's use case, use mirror representations.

The Interplay between Movements and Representations

Although the differences between the representations are significant (see Table 1), more interesting and context-dependent findings were unveiled when analyzing the effect of the representations separately for each movement type. This allowed us to extend the recommendations to consider not only recognition time and accuracy, but other aspects such as privacy concerns, social embarrassment, playfulness and interaction times. It is important to study and understand the physical characteristics of the public display's deployment space, and identify how they affect passersby behavior [20]. Hence, the following recommendations inform practitioners on how to design user representations, given particular setups that allow for a specific type of movement.

NoMove: Standing Still

Deployments in which passersby are expected to perceive the display only when they are stationary (e.g., in an elevator), abstract objects and skeletons should be avoided as they result in significantly longer recognition times and more errors. Mirrors perform significantly better than all other representations in NoMove. Silhouettes perform significantly better than abstract objects and skeletons and its performance is acceptable (see Figures 5 and 6).

Recommendation 2: If users are expected to perceive the display while stationary, silhouettes are acceptable and come in second place after mirrors. Abstract objects and skeletons should be avoided.

Arms: Moving Arms While Stationary

If users are able to move their arms but not walk around (e.g., while standing in a queue), skeletons, silhouettes and mirrors reach their peak performance (see Figure 5).

Recommendation 3: Multiuser systems should encourage arm movements (e.g., interaction via mid-air gestures) as they positively influence recognition time and accuracy.

Abstract objects also perform significantly worse than other representations in the Arms condition. While mirrors perform significantly better than both of them, the efficiency of skeletons and silhouettes is almost similar in that condition. Hence a designer's choice of silhouettes or skeletons when the Arms condition is expected should not rely on these factors, since they perform almost equally efficient in terms of recognition time and accuracy.

Recommendation 4: In contexts where users are expected to move their arms while stationary, use skeletons for playful and longer interaction times, or silhouettes for serious interactions (e.g., task-oriented interactions [23]). The differences between them are negligible in terms of recognition time and accuracy but both come next after mirrors. Do not use abstract objects when users are expected to move their arms.

Walk: Walking Without Moving Arms

In the case of deployments where users can walk but cannot move their arms (e.g., both hands are occupied by holding a cup of coffee and a bag, or a suitcase and a flight ticket), differences in performance between the different representations are almost negligible. Although mirrors are significantly faster to recognize in comparison to abstract objects, skeletons and silhouettes, the three representations perform almost equally good. This means that if one of the latter three is to be used, design decisions should base their choice of representation on factors other than recognition accuracy and time.

Recommendation 5: Although mirrors outperform them, the differences between abstract objects, skeletons, and silhouettes in recognition time and accuracy are negligible when walking without moving the arms. Use abstract objects when a top-view of passersby is desired, skeletons for playful and longer interaction times, and silhouettes for serious and task-oriented interactions.

Walk+Arms: Walking and moving arms

By walking and moving the arms (e.g., unrestricted use in museums or while shopping) abstract objects perform the worst among all representations. As previously mentioned, it is expected that moving the arms distracts participants, and could induce a cognitive load [27].

Recommendation 6: If participants are walking and moving their arms, do not use abstract objects, but rather mirrors, silhouettes or skeletons.

FUTURE WORK

Previous work as well as our results show that skeletons result in longer interaction times due to their playfulness. Meanwhile, our results show that mirrors are always faster to recognize compared to skeletons. Hence, a promising setup would be to attract the user's attention first using mirrors to increase the chances the passersby recognize them, then switch to skeletons as soon as the participant starts interaction, to increase interaction times. Previous work has shown that switching from a user representation to another does not necessarily confuse the user [35]. Further experimentation in this direction could result in an optimal sequence that utilizes a different representation at each interaction stage [21].

Another interesting follow-up work would be to experiment with simultaneously showing different representations. Systems can decide which representation to use for each user based on the user's movement type; standing users (NoMove) would see their mirror or silhouette reflection, whilst walking ones (Walk and Walk+Arms) would see a skeleton.

Gaze-enabled public displays are becoming more common as gaze shows promise in addressing many challenges of public displays (see [12] for an overview). One direction for future work is to show the representation right where the user is looking. Still, users would have to distinguish their representations in case several users are looking at the same area, but the probability of this happening decreases in larger displays.

Finally, it would be interesting to conduct follow up studies with different groups of participants, and with different space constraints.

CONCLUSION

In this work, we studied how well passersby can distinguish their own representations from those of others on a large public display. We identified 5 main strategies that users employ to identify themselves in a pre-study. In a follow-up study we quantified the time and accuracy of recognizing one's own user representation. Furthermore we introduced 6 recommendations to help designers choose the suitable user representation depending on the context of their deployments.

ACKNOWLEDGEMENTS

This work was funded, in part, by the Cluster of Excellence on Multimodal Computing and Interaction (MMCI) at Saarland University, Germany, and by the Bavarian State Ministry of Education, Science and the Arts in the framework of the Center Digitization.Bavaria (ZD.B). This research was supported by the German Research Foundation (DFG), Grant No. AL 1899/2-1. We would like to thank Carina Hoppenz for her help with the statistical analysis.

REFERENCES

1. Christopher Ackad, Andrew Clayphan, Martin Tomitsch, and Judy Kay. 2015. An In-the-wild Study of Learning Mid-air Gestures to Browse Hierarchical Information at a Large Interactive Public Display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 1227–1238. DOI: http://dx.doi.org/10.1145/2750858.2807532
2. Christopher Ackad, Martin Tomitsch, and Judy Kay. 2016. Skeletons and Silhouettes: Comparing User Representations at a Gesture-based Large Display. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2343–2347. DOI: http://dx.doi.org/10.1145/2858036.2858427
3. Christopher Ackad, Rainer Wasinger, Richard Gluga, Judy Kay, and Martin Tomitsch. 2013. Measuring Interactivity at an Interactive Public Information Display. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration (OzCHI '13)*. ACM, New York, NY, USA, 329–332. DOI: http://dx.doi.org/10.1145/2541016.2541091
4. Moritz Behrens, Ava Fatah gen. Schieck, Efsathia Kostopoulou, Steve North, Wallis Motta, Lei Ye, and Holger Schnadelbach. 2013. Exploring the Effect of Spatial Layout on Mediated Urban Interactions. In *Proceedings of the 2Nd ACM International Symposium on Pervasive Displays (PerDis '13)*. ACM, New York, NY, USA, 79–84. DOI: http://dx.doi.org/10.1145/2491568.2491586
5. Victor Cheung and Stacey Scott. 2016. Proxemics-Based Visual Concepts to Attract and Engage Public Display

Users: Adaptive Content Motion and Adaptive User Shadow. In *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces (ISS '16)*. ACM, New York, NY, USA, 473–476. DOI: http://dx.doi.org/10.1145/2992154.2996875

6. James E. Cutting and Lynn T. Kozlowski. 1977. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society* 9, 5 (1977), 353–356. DOI: http://dx.doi.org/10.3758/BF03337021

7. Nigel Davies, Sarah Clinch, and Florian Alt. 2014. *Pervasive Displays: Understanding the Future of Digital Signage* (1st ed.). Morgan & Claypool Publishers.

8. Ivan Elhart, Federico Scacchi, Evangelos Niforatos, and Marc Langheinrich. 2015. ShadowTouch: A Multi-user Application Selection Interface for Interactive Public Displays. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM, New York, NY, USA, 209–216. DOI: http://dx.doi.org/10.1145/2757710.2757735

9. Jonna Häkkilä, Olli Koskenranta, Maaret Posti, Leena Ventä-Olkkonen, and Ashley Colley. 2013. Clearing the Virtual Window: Connecting Two Locations with Interactive Public Displays. In *Proceedings of the 2Nd ACM International Symposium on Pervasive Displays (PerDis '13)*. ACM, New York, NY, USA, 85–90. DOI: http://dx.doi.org/10.1145/2491568.2491587

10. Giulio Jacucci, Ann Morrison, Gabriela T. Richard, Jari Kleimola, Peter Peltonen, Lorenza Parisi, and Toni Laitinen. 2010. Worlds of Information: Designing for Engagement at a Public Multi-touch Display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2267–2276. DOI: http://dx.doi.org/10.1145/1753326.1753669

11. Mohamed Khamis, Florian Alt, and Andreas Bulling. 2015. A Field Study on Spontaneous Gaze-based Interaction with a Public Display Using Pursuits. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*. ACM, New York, NY, USA, 863–872. DOI: http://dx.doi.org/10.1145/2800835.2804335

12. Mohamed Khamis, Florian Alt, and Andreas Bulling. 2016. Challenges and Design Space of Gaze-enabled Public Displays. In *Ajunct Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 10. DOI: http://dx.doi.org/10.1145/2968219.2968342

13. Mohamed Khamis, Daniel Buschek, Tobias Thieron, Florian Alt, and Andreas Bulling. 2017a. EyePACT: Eye-Based Parallax Correction on Touch-Enabled Interactive Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 146 (Dec. 2017), 18 pages. DOI: http://dx.doi.org/10.1145/3161168

14. Mohamed Khamis, Regina Hasholzner, Andreas Bulling, and Florian Alt. 2017b. GTmoPass: Two-factor Authentication on Public Displays Using Gaze-touch Passwords and Personal Mobile Devices. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays (PerDis '17)*. ACM, New York, NY, USA, Article 8, 9 pages. DOI: http://dx.doi.org/10.1145/3078810.3078815

15. David Kieras. 2001. Using the keystroke-level model to estimate execution times. *University of Michigan* (2001).

16. Christian Lander, Marco Speicher, Denise Paradowski, Norine Coenen, Sebastian Biewer, and Antonio Krüger. 2015. Collaborative Newspaper: Exploring an Adaptive Scrolling Algorithm in a Multi-user Reading Scenario. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM, New York, NY, USA, 163–169. DOI: http://dx.doi.org/10.1145/2757710.2757734

17. Eva Lösch, Florian Alt, and Michael Koch. 2017. Mirror, Mirror on the Wall: Attracting Passers-by to Public Touch Displays With User Representations. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, New York, NY, USA, 22–31. DOI: http://dx.doi.org/10.1145/3132272.3134129

18. Andrés Lucero, Jussi Holopainen, and Tero Jokela. 2012. MobiComics: Collaborative Use of Mobile Phones and Large Displays for Public Expression. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '12)*. ACM, New York, NY, USA, 383–392. DOI:http://dx.doi.org/10.1145/2371574.2371634

19. Ville Mäkelä, Tomi Heimonen, and Markku Turunen. 2014. Magnetic Cursor: Improving Target Selection in Freehand Pointing Interfaces. In *Proceedings of The International Symposium on Pervasive Displays (PerDis '14)*. ACM, New York, NY, USA, Article 112, 6 pages. DOI:http://dx.doi.org/10.1145/2611009.2611025

20. Ville Mäkelä, Sumita Sharma, Jaakko Hakulinen, Tomi Heimonen, and Markku Turunen. 2017. Challenges in Public Display Deployments: A Taxonomy of External Factors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3426–3475. DOI: http://dx.doi.org/10.1145/3025453.3025798

21. Daniel Michelis and Jörg Müller. 2011. The Audience Funnel: Observations of Gesture Based Interaction With Multiple Large Displays in a City Center. *International Journal of HumanâĂŞComputer Interaction* 27, 6 (2011), 562–579. DOI: http://dx.doi.org/10.1080/10447318.2011.555299

22. Robert W. Mitchell. 1993. Mental models of mirror-self-recognition: Two theories. *New Ideas in Psychology* 11, 3 (1993), 295 – 325. DOI: http://dx.doi.org/10.1016/0732-118X(93)90002-U

23. Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. 2010. Requirements and Design Space for Interactive Public Displays. In *Proceedings of the 18th ACM International Conference on Multimedia (MM '10)*. ACM, New York, NY, USA, 1285–1294. DOI: http://dx.doi.org/10.1145/1873951.1874203

24. Jörg Müller, Gilles Bailly, Thor Bossuyt, and Niklas Hillgren. 2014a. MirrorTouch: Combining Touch and Mid-air Gestures for Public Displays. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 319–328. DOI: http://dx.doi.org/10.1145/2628363.2628379

25. Jörg Müller, Dieter Eberle, and Konrad Tollmar. 2014b. Communiplay: A Field Study of a Public Display Mediaspace. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1415–1424. DOI: http://dx.doi.org/10.1145/2556288.2557001

26. Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. 2012. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 297–306. DOI: http://dx.doi.org/10.1145/2207676.2207718

27. Philipp Panhey, Tanja Döring, Stefan Schneegass, Dirk Wenig, and Florian Alt. 2015. What People Really Remember: Understanding Cognitive Effects When Interacting with Large Displays. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces (ITS '15)*. ACM, New York, NY, USA, 103–106. DOI:http://dx.doi.org/10.1145/2817721.2817732

28. Gonzalo Parra, Robin De Croon, Joris Klerkx, and Erik Duval. 2014. Quantifying the Interaction Stages of a Public Display Campaign in the Wild. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14)*. ACM, New York, NY, USA, 757–760. DOI: http://dx.doi.org/10.1145/2639189.2639216

29. Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. 2008. It's Mine, Don'T Touch!: Interactions at a Large Multi-touch Display in a City Centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 1285–1294. DOI: http://dx.doi.org/10.1145/1357054.1357255

30. Martin Tomitsch, Christopher Ackad, Oliver Dawson, Luke Hespanhol, and Judy Kay. 2014. Who Cares About the Content? An Analysis of Playful Behaviour at a Public Display. In *Proceedings of The International Symposium on Pervasive Displays (PerDis '14)*. ACM, New York, NY, USA, Article 160, 6 pages. DOI: http://dx.doi.org/10.1145/2611009.2611016

31. Nina Valkanova, Robert Walter, Andrew Vande Moere, and Jörg Müller. 2014. MyPosition: Sparking Civic Discourse by a Public Interactive Poll Visualization. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1323–1332. DOI:http://dx.doi.org/10.1145/2531602.2531639

32. Jouni Vepsäläinen, Petri Savolainen, Jouni Ojala, Antonella Di Rienzo, Matti Nelimarkka, Kai Kuikkaniemi, Sasu Tarkoma, and Giulio Jacucci. 2016. Web-Based Public-Screen Gaming: Insights from Deployments. *IEEE Pervasive Computing* 15, 3 (July 2016), 40–46. DOI: http://dx.doi.org/10.1109/MPRV.2016.60

33. Daniel Vogel and Ravin Balakrishnan. 2004. Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, New York, NY, USA, 137–146. DOI: http://dx.doi.org/10.1145/1029632.1029656

34. Robert Walter, Gilles Bailly, and Jörg Müller. 2013. StrikeAPose: Revealing Mid-air Gestures on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 841–850. DOI: http://dx.doi.org/10.1145/2470654.2470774

35. Robert Walter, Gilles Bailly, Nina Valkanova, and Jörg Müller. 2014. Cuenesics: Using Mid-air Gestures to Select Items on Interactive Public Displays. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 299–308. DOI: http://dx.doi.org/10.1145/2628363.2628368

36. Robert Walter, Andreas Bulling, David Lindlbauer, Martin Schuessler, and Jörg Müller. 2015. Analyzing Visual Attention During Whole Body Interaction with Public Displays. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 1263–1267. DOI: http://dx.doi.org/10.1145/2750858.2804255

37. Niels Wouters, John Downs, Mitchell Harrop, Travis Cox, Eduardo Oliveira, Sarah Webber, Frank Vetere, and Andrew Vande Moere. 2016. Uncovering the Honeypot Effect: How Audiences Engage with Public Interactive Systems. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 5–16. DOI: http://dx.doi.org/10.1145/2901790.2901796

38. Yanxia Zhang, Ken Pfeuffer, Ming Ki Chong, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2017. Look Together: Using Gaze for Assisting Co-located Collaborative Search. *Personal Ubiquitous Comput.* 21, 1 (Feb. 2017), 173–186. DOI: http://dx.doi.org/10.1007/s00779-016-0969-x